International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Corporate Bankruptcy Prediction: A Case of Emerging Economies

Suresh Ramakrishnan[a], Maryam Mirzaei[b]*, Muhammad Naveed[c]

[a,b,c] *Department of Finance, Faculty of Management, Universiti Teknologi Malaysia, Johor, Malaysia*

[b]*Email: mmirzai72@yahoo.ca*

Abstract

Bankruptcy has recently upraised as an excessive concern due to the recent world crisis. Early forecasting of firms bankruptcy provides decision-support information for financial and regulatory institutions. In spite of several progressive methods that have widely been proposed, this area of research is not out dated and still needs further examination. In this paper, the performance of different multiple classifier systems are assessed in terms of their capability to appropriately classify bankruptcy and non-bankruptcy Iranian firms listed in Tehran Stock Exchange (TSE). On the other hand, TSE have had very high return which provided more than 140 percent return in last year. For this reason, TSE could be more attractive for investors. Most data mining techniques provided significant improvements over the linear regression. In addition, non-linear classifiers afford enhancement in performance over the linear techniques.

Keywords: Bankruptcy prediction; Financial distress; Machine learning

1. Introduction

Due to the significant consequences which bankruptcy imposes on different groups of society as well the noteworthy troubles qualified by firms during the Global Financial Crisis, the crucial importance of measuring and providing for credit risk have highlighted.

---

* Corresponding author.

E-mail address: mmirzai72@yahoo.ca.

Since the mid-1990s, there has been growing concern in emerging and developing economies among researchers. One of the least studied emerging markets is the Tehran Stock Exchange (TSE), however a study of the TSE would contribute to the literature on emerging and developing markets finance especially Middle East. The value of Tehran Stock Exchange return was increased by 140 percent at the end of 2013. Regarding the growth in financial services, there have been swelling sufferers from off ending loans. Therefore, bankruptcy risk forecasting is a critical part of a financial institution's loan approval decision processes.

Bankruptcy risk prediction is a procedure that determines how likely applicants are to bankruptcy with their repayments. Review of literature on the subject confirmed hand full of studies conducted in the last four decades. Despite of these studies, the recent credit crisis indicated that yet there are areas of the study that needs researchers' attention. Moreover, emerging of the regulatory changes such as Basel III accord and the need for more precise and comprehensive risk management procedures justifies need of research in area of credit risk modeling and banking supervision. This requirement like these pushes companies especially banks and insurance companies to have a very robust and transparent risk management system.

Since the study of [15], bankruptcy prediction becomes a challenging issue in corporate finance. Earlier, most of the studies on bankruptcy risk focused on firm-specific indicators as a predictor of firms bankruptcy across United States including [24, 26]. Although, majority of the studies used the firm-specific variables, some researchers tried to use some other indicators such as interest rate, stock index return and GDP that affects bankruptcy prediction. As a result of relationship between general economic and bankruptcy rates, some attempts have been made to predict bankruptcy based on macroeconomic variables. Earlier surveys using U.S. firms [22, 34] revealed that macroeconomic indicators affect bankruptcy prediction.

The majority of the discussion related to bankruptcy prediction develops around the decisive works of [2, 28, 35, 32]. The author of [2] applied Multivariate Discriminant Analysis (MDA) for the first time to classify failed and non-failed U.S firms. Researchers still use this model as a benchmark to predict firm bankruptcy. Altman's Z-score model is a linear analysis of five ratios and this score is a basis for firm classification. Besides, [8] employed the same MDA technique for bankruptcy prediction some years prior to failure. Similarly, to assess the predictive accuracy of accounting ratios, [21] measured the prediction achievement of a selected set of accounting ratios for U.S ventures. The use of technique as a benchmark tool for bankruptcy prediction shows researchers trust on the technique [19].

Risk bankruptcy prediction is conventionally observed from a binary classification standpoint. Therefore, classification or regression methods used to generate a classifier which creates a numerical output to demonstrate the probability of a firm to repay its obligations (good applicant) or bankruptcy or display undesirable behavior (bad applicant). Adding to the debate on bankruptcy prediction methodology Crook et al, (2007) suggested logistic regression as one of the leading methods. Similarly, linear regression/ discriminant analysis and decision trees are other popular methods [11, 9]. These methods are easy to develop and levels of performance are acceptable. In addition, there is abundance of literature revolving around the application of other methodologies. Some includes statistical techniques such as survival analysis, graphical models and Markov chains. Other methods are from the machine learning/data mining approaches such as; support vector machines [5] and genetic algorithms. Principally, machine learning techniques have shown superior performance over statistical ones [29]. Putting simply, nonlinear approaches such as neural network and support vector machines outperform other methods but by a small margin [10] in order to non-linearity dependence between some financial ratios and bankruptcy probability. There are no hard and fast rules as to which methodology; a model developer should approve for this improbability. Consequently, a promising approach to adopt this uncertainly is to construct a number of classifiers using various techniques, and then choose one that proves best against the problem. It has been established that combining a set of independent classifiers with adequate accuracy leads to better performance, provided that the diversity among accurate based classifiers in an ensemble system is enforced in some way. Although, the research into classifier combination is remarkable, a few report empirical findings have been done on corporate bankruptcy. In addition, there has been little effort to compare the wide range of classifiers within previous studies.

Significant advances have been made in the past few decades regarding methodologies for bankruptcy prediction. [6] introduced the Naïve Bayes approach using a single variable and Altman in 1968 suggested the use of Linear Discriminant Analysis (LDA). Since then several contributions have been made to improve the Altman's results, using different techniques. The use of data mining techniques such as Artificial Neural Networks (ANN), decision trees, and Support Vector Machine (SVM) for bankruptcy prediction started in the late 1980s [31].

The authors in [16] used Decision Trees first time for bankruptcy prediction. Using this model, they classified firms to failed and non-failed based on firm-level and country-level factors. According to their results, this technique allows for an easy identification of the most significant characteristics in bankruptcy prediction. In another study, Quinlan (1986) noted that decision trees method can deal with noise or non-systematic errors in the values of features. There are some other studies which predicted bankruptcy using this method such as [25]. Detailed examination of corporate bankruptcy prediction by [30] showed a better performance of the hybrid model. They used four different techniques to predict corporate bankruptcy, which two of the methods were statistical and the outstanding two models were machine learning techniques. In different but related work, the authors in [30] suggested a model using genetic algorithms technique. Some other related studies have employed Artificial Neural Networks to predict bankruptcy.

Artificial Neural Networks was first demonstrated experimentally by [20] to analyze bankrupt companies. Since then the method became a common accuracy amongst. Recently, some of the main commercial loan bankruptcy prediction products applied ANN technique. For example, Moody's public firm risk model ANN and many banks and financial institutions have developed this method for bankruptcy prediction [4]. More recently, the support vector machine was commenced for bankruptcy risk investigation. This technique which is based on statistical learning theory compared with the traditional methods is more accurate in predicting bankruptcy likelihood [18].

2. Methodology

2.1. Logistic Regression

Logistic regression is a type of regression methods [1] where the dependent variable is discrete or categorical, for instance, bankruptcy (1) and non-bankruptcy (0). Logistic regression examines the effect of multiple independent variables to forecast the association between them and dependent variable categories. According to [27,23] was the first researcher who used logistic technique in corporate bankruptcy perspective. He employed this technique to examine failures in the U.S. banking sector. Subsequently, [28] applied logistic regression more generally to a sample of 105 bankrupt firm and 2,000 non-bankrupt companies. His model did not discriminate between failed and non-failed companies as well as the multiple discriminant analysis (MDA) models reported in previous studies. According to [13], logistic regression is in the second place, after MDA, in bankruptcy prediction models.

2.2. Decision Tree

Decision trees are the most popular and powerful techniques for classification and prediction. The foremost cause behind their recognition is their simplicity and transparency, and consequently relative improvement in terms of interpretability. Decision tree is a non-parametric and introductory technique, which is capable to learn from examples by a procedure of simplification. [16] first time employed decision trees to forecast bankruptcy. Soon after, some researchers applied this technique to predict bankruptcy and bankruptcy including [17, 25].

2.3. Neural Networks

Neural networks (NNs), usually non-parametric techniques have been used for a variety of classification and regression problems. They are characterized by associates among a very large number of simple computing processors or elements (neurons). Corporate bankruptcy have predicted using neural networks in early 1990s and since then more researchers have used this model to predict bankruptcy. As a result, there are some main profitable loan bankruptcy prediction products which are based on neural network models. Also, there are different evidence from many banks which have already expanded or in the procedure of developing bankruptcy prediction models using neural network [4]. This technique is flexible to the data characteristics and can deal with different non-linear functions and parameters also compound prototypes. Therefore, neural networks have the ability to deal with missing or incomplete data [33].

2.4. Support Vector Machines

Among different classification techniques, Support Vector Machines are considered as the best classification tools accessible nowadays. There are a number of empirical results attained on a diversity of classification (and regression) tasks complement the highly appreciated theoretical properties of SVMs. A support vector machine (SVM) produces a binary classifier, the so-called optimal separating hyper planes, through extremely nonlinear mapping the input vectors into the high-dimensional feature space. SVM constructs linear model to estimate the decision function using non-linear class boundaries based on support vectors.Support vector machine is based on a linear model with a kernel function to implement non-linear class boundaries by mapping input vectors non-linearly into a high-dimensional feature space.

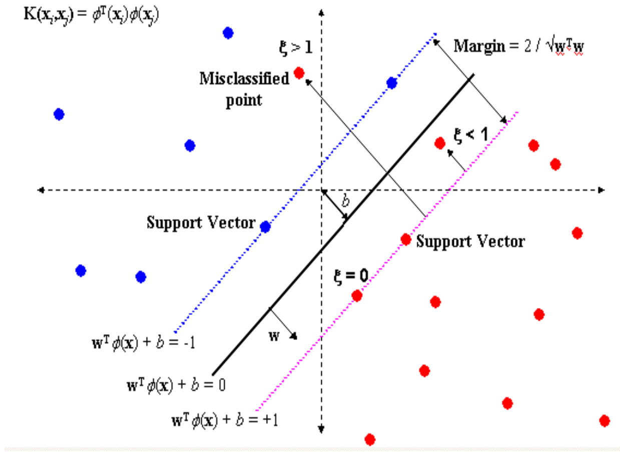

Figure 1: The SVM learns a hyperplane which best separates the two classes.

Based on conceptual elements of statistical learning and the potential of SVMs for firm rating, for the linear classification problem a SVM is defined and this method is simplified for nonlinear cases. In the linear case (figure 1) the following inequalities hold for all n points of the training set:

$$\text{Min}$$

$$x_i^T w+b \geq 1-\xi_i \qquad \text{for} \qquad y_i=1,$$

$$x_i^T w+b \leq -1+\xi_i \qquad \text{for} \qquad y_i=-1,$$

$$\xi_i \geq 0$$

This can be combined into two constraints:

$$Y_i\,(x_i^T w+b) \geq 1-\xi_i$$

$$\xi_i \geq 0$$

The basic idea of the SVM classification is to find such a separating hyperplane that corresponds to the largest possible margin between the points of different classes.

3. Empirical experiment

3.1. Data Description

The dataset was used to classify a set of firms into those that would bankruptcy and those that would not bankruptcy on loan payments. It consists of 217 observations of Iranian companies. All of them were or still are listed on the Tehran Stock Exchange (TSE). Of the 217cases for training,100 belong to the bankruptcy case under paragraph 141 of Iran Trade Law and the other 117to non-bankruptcy case.

The 21 significant variables in this study were selected by using a two stages predictive variable selection process. At the first stage, bankruptcy prediction literature was reviewed and 65 variables from more than 230 financial ratios were selected as predictive variables. These financial ratios were chosen based on their popularity in the literature. In the second stage, 21 variables were selected based on the availability of the necessary data. The components of the financial ratios which are estimated from data are explained below and table 1 shows the summary statistics for selected variables for bankruptcy and non-bankruptcy firms.

To select the variables, two approaches including linear regression and decision tree analysis were used. The most significant variables based on two methods were identified. These variables selected from the 21 indicators for the model which could best discriminate the bankruptcy firms from the non-bankruptcy firms. These selected financial ratios include: EBIT to total assets (X1), current assets to total assets (X5), net profit to liability (X11), working capital to total assets (X6) and net profit to sale (X16).

3.2. Experimental Results

This section demonstrates the results and main findings of the analysis in order to probability of default. Considerable attention has been devoted to financial ratio analysis for classifying failed and non-failed companies. Based on the results, the most important indicators are liquidity and profitability. The liquidity position of a firm is when the firm is able to meet certain financial obligations. In other words, the firm has the ability to repay its obligations without incurring too much cost. Based on the results liquidity is one of the significant indicators which affects the probability of default. Beside this, more profitable firms are less probable to face default. The significant indicators then have been employed to model probability of default using machine learning techniques including, decision tree, neural network and support vector machine.

Comparison of forecasting accuracy reveals that the SVM has a lower model risk than other models. According to the results, SVM is the best. The performance of logistic regression is significantly worse than other approaches. Generally, the findings for the classifiers are not predominantly unexpected and are well-matched with previous empirical researches of classifier performance for default risk data sets especially in case of SVM classifier. SVM with a high generalization capacity seems to be a capable technique for default prediction in Iran as an emerging economy. Also, table 2 shows the performance accuracy of different models.

Roc curve plots the type II error against one minus the type I error. In the case of default prediction in this study, it describes the percentage of non-defaulting firms that must be inadvertently denied credit (Type II) in order to avoid lending to a specific percentage of defaulting firms (1- Type I) when using a specific model. Figure 2, shows the ROC curve for different classifiers. The results also state the improvement by the machine learning techniques is significant.

5. Conclusion

Default prediction takes an important role in the prevention of corporate default, which makes the accuracy of default prediction model be widely concerned by researchers. Appropriate identification of firms 'approaching default is undeniably required.

Table 1: The summary statistics for selected variables for bankruptcy and non-bankruptcy firms

	Definition of variable	Means of non-bankruptcy companies	Means of bankruptcy companies	Test of equality of group means		Definition of variable	Means of non-bankruptcy companies	Means of bankruptcy companies	Test of equality of group means
1	EBIT/TA	0.155647	-0.02608	0	12	NP/E	-0.08432	-1.0931	0.079
2	Ca/TA	0.086677	0.031281	0	13	S/TA	2.611479	2.424024	0.499
3	Ca/CL	1.854502	1.178482	0	14	S/CA	4.410169	4.135828	0.511
4	CA/CL	0.618186	0.584403	0.271	15	R/L	0.807465	0.561762	0
5	CA/TA	0.151896	-0.05514	0	16	NP/S	0.039157	-0.04186	0
6	WC/TA	0.098689	-0.87107	0.174	17	L/TA	0.522716	0.704322	0
7	WC/S	9.504097	16.74657	0.147	18	L/E	16.12618	7.795392	0.483
8	S/R	0.095607	-0.06198	0	19	LL/E	6.580743	0.176757	0.315
9	NP/TA	0.095607	-0.06197	0	20	CL/E	9.545434	7.618636	0.738
10	GP/S	0.602299	-0.616	0	21	CA/S	0.933685	0.426074	0.441
11	NP/L	0.347821	0.080305	0					
	EBIT:	Earning Before Interests and Taxes				GP:	Gross Profit		
	TA:	Total assets				L:	Liabilities		
	Ca:	Cash				TI:	Total Income		
	CL:	Current Liabilities				LL:	Long Term Liabilities		
	CA:	Current Assets				E:	Equity		
	WC:	Working Capital				R:	Receivables		
	S:	Sale				NP:	Net Profit		

Table 2: Performance of classifier systems

	Classifier system	% Accuracy	% Misclassified	% ROC Area
Baseline models	LR	69.12	30.76	82.3
	NN	83.41	17.38	91.3
	DT	74.79	19.12	83.8
	SVM	84.79	14.42	92.5

By this time, various methods have been used for predicting default. The use of machine learning techniques is becoming common in different studies. Since the pioneering works of [6] and [2], many researches have been developed to predict corporate default using financial ratios, and it seems that these might be other quantitative and qualitative variables that can help prediction. The results of this study states liquidity of the firm and profitability have proved significant effect on probability of default. In addition, machine learning classifiers outperform the statistical methods and among them SVM shows higher accuracy more than other three methods with statistical significance and especially suits for Iranian listed companies. Therefore, this study contributes to provide incremental evidence for default prediction research based on machine learning and guide the real world practice of DP to some extent. However, this study also has the limitation that the experimental data sets are only collected from Iranian listed companies, and further investigation can be done based on other countries' real world data sets in future study.

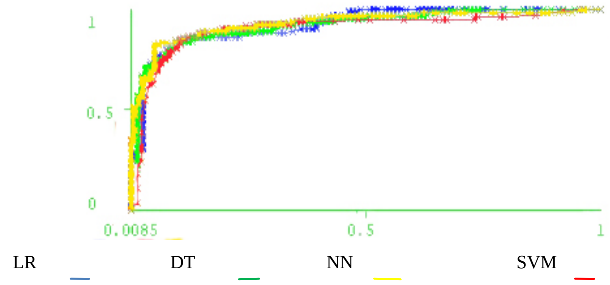

Fig 2. ROC curve

References

[1] P.D.Allison, (2001). Logistic Regression Using the SAS System: Theory and Application. *Cary, NC: SAS Publishing, BBU Press.*

[2] Edward I.Altman, (1968). Financial Ratios, Discriminant Analaysis and the Prediction of Corporate Bankruptcy. *Journal of Finance, 23*(4), 589-609.

[4] Atiya, A. (2001). Bankruptcy Prediction for Credit Risk using Neural Networks: A Survey and New Results. *IEEE Transactions on Neural Networks, 12*, 929-935.

[5] Baesens, B., Gestel, T.V., Viaene, S., Stepanova, M., Suykens, J., Vanthienen, J., 2003a.Benchmarking state-of-the-art classification algorithms for credit scoring.*Journal of the Operational Research Society* 54, 627–635.

[6] Beaver, William H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research, 4*(3), 71-111.

[7] Blum, James D. (1977). Financial Statement Analysis: A New Approach. *Journal of Finance*, *32*(3), 958-961.

[8] Blum, Marc. (1974). Failing Company Discriminant Analysis. *Journal of Accounting Research, 12*(1), 1-25.

[9] Boyle, M., Crook, J. N., Hamilton, R., Thomas, L. C. 1992. Methods for credit scoring appliedto slow payers. In Credit Scoring and Credit Control, Thomas, L. C., Crook, J. N.,Edelman, D. B., eds., Oxford University Press, Oxford, 75-90.

[10] Crook, J., Edelman D., Thomas, L. 2007. Recent developments in consumer credit risk assessment *European Journal of Operational Research*, 183 (3): 1447- 1465.

[11] Durand, D. (1941). Risk elements in consumer instalment financing. (Technical edition) ByDavid Durand. National bureau of economic research [New York].

[12] Deakin, Edward B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research, 10*(1), 167-179.

[13] Dimitras, A. I., Zanakis, S. H., & Zopounidis, C. (1996). A survey ofbusiness failure with an emphasis onprediction methods andindustrialapplication. *European Journal of OperationalResearch*, 90, 487–513.

[14] Hennawy, R. H. A., & Morris, R. C. (1983). THE SIGNIFICANCE-OF BASE YEAR IN DEVELOPING FAILURE PREDICTION MODELS. *Journal of Business Finance & Accounting, 10*(2), 209-223.

[15] FitzPatrick, Paul Joseph. (1932). *A comparison of the ratios of successful industrial enterprises with those of failed companies*: Washington.

[16] Frydman, H., Altman, E., & Kao, D. (1985). Introducing recursive partitioning for financial classification: the case of financial distress. *Journal of Finance*, 269-291.

[17] Gepp, Adrian, Kumar, Kuldeep, & Bhattacharya, Sukanto. (2010). Business failure prediction using decision trees. *Journal of Forecasting, 29*(6), 536-555. doi: 10.1002/for.1153

[18] Härdle, Wolfgang, Moro, Rouslan, & Schäfer, Dorothea. (2005). Predicting Bankruptcy with Support Vector Machines. *Statistical Tools in Finance & Insurance*, 225-248.

[19] Hillegeist, S.A., Keating, E.K., Cram, D.P. and Lundstedt, K.G. (2004). Assessing the Probability of Bankruptcy, Review of Accounting Studies, 9(1), 5–34.

[20] Hertz, J., Krogh, A., & Palmer, R.G. (1991). The Theory of Neural Network Computation. *Addison Welsey: Redwood, CA.*

[21] Libby, Robert. (1975). Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence. *Journal of Accounting Research*, *13*(1), 150-161.

[22] Liou, Dah-Kwei, & Smith, Malcolm. (2007). Macroeconomic Variables and Financial Distress. *Journal of Accounting, Business & Management, 14,* 17-31.

[23] Martin, D. (1977). Early warnings of bank failure: A logit regression approach. *Journal of Banking and Finance*, *1*, 249-276.

[24] Merwin, C. (1942). Financing small corporations in five manufB.cturing industries, 1926-1936. *New York: National Bureau of Economic Research*.

[25] Messier, JR. W., & Hansen, J. (1988). Inducing rules for expert system development: an example using bankruptcy and bankruptcy data. *Management Science, 34,* 1403-1415.

[26] Meyer, Paul A., & Pifer, Howard W. (1970). Prediction of bank failures. *Journal of Finance*, *25*(4), 853-868.

[27] Morris, Richard. (1997). *Early warning indicators of corporate failure : a critical review of previous research and further mpirical evidence*. Ashgate.

[28] Ohlson, James A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, *18*(1).

[29] Ravi Kumar, P., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review. *European Journal of Operational Research*, *180*(1), 1-28.

[30] Shin, K.S., and Lee, Y.J. (2002). A Genetic Algorithm Application in Bankruptcy Prediction Modeling. Expert Systems with Applications, 9, 503-512.

[31] Shin, K.S., lee, T.S., & Kim, H.J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert system Application, 28*(1), 127-135.

[32] Shumway, Tyler. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business*, *74*(1), 101-124.

[33] Smith, C.W., & Stulz, R.M. (1985). The Determinants of Firms' Hedging Policies. *Journal of Financial and Quantitative Analysis, 20*(4), 391-405.

[34] Vassalou, Maria, & Yuhang, Xing. (2004). Bankruptcy Risk in Equity Returns. *Journal of Finance, 59*(2), 831-868. doi: 10.1111/j.1540-6261.2004.00650.x

[35] Zmijewski, M.E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research* 22, 59–86.